Wave data prediction and reconstruction by recurrent neural networks at the nearshore area of Norderney

Christoph Jörges[1], Cordula Berkenbrink[2] and Britta Stumpe[1]

1 Department of Human-Environment Research, Institute of Geography, Bergische Universität Wuppertal, Germany
2 Coastal Research Station Norderney, NLWKN, Germany

EGU GENERAL ASSEMBLY 2020, ONLINE

(1) Introduction

Sea level rise, a possible increase in frequency and intensity of storms and other effects of global warming exert pressure on the coastal regions of the North Sea. As well as for building coastal protection or offshore structures, detailed knowledge of wave data, especially the wave height, is of particular interest. Therefore, the wave climate at the island Norderney is measured by buoys since the early 1990s. Caused by crossing ships or weather impacts, these buoys can be damaged, leading to a huge amount of missing data in the time series, which are the basis for numerical modelling, statistical analysis and developing coastal protection.

ANNs are a common method to forecast and reconstruct wave heights nowadays. This study shows a new approach for significant wave height data measured by buoys in the nearshore area of the Norderney coastline. Buoy data of the period 2004 to 2017 from the NLWKN - Coastal Research Station at Norderney were used to train three different statistical and machine learning models namely linear regression, feed-forward neural network and long short-term memory (LSTM), respectively. Besides calculated sea state parameter, an energy density spectrum is being tested as predictor.

(2) Data and Methods

Research Area

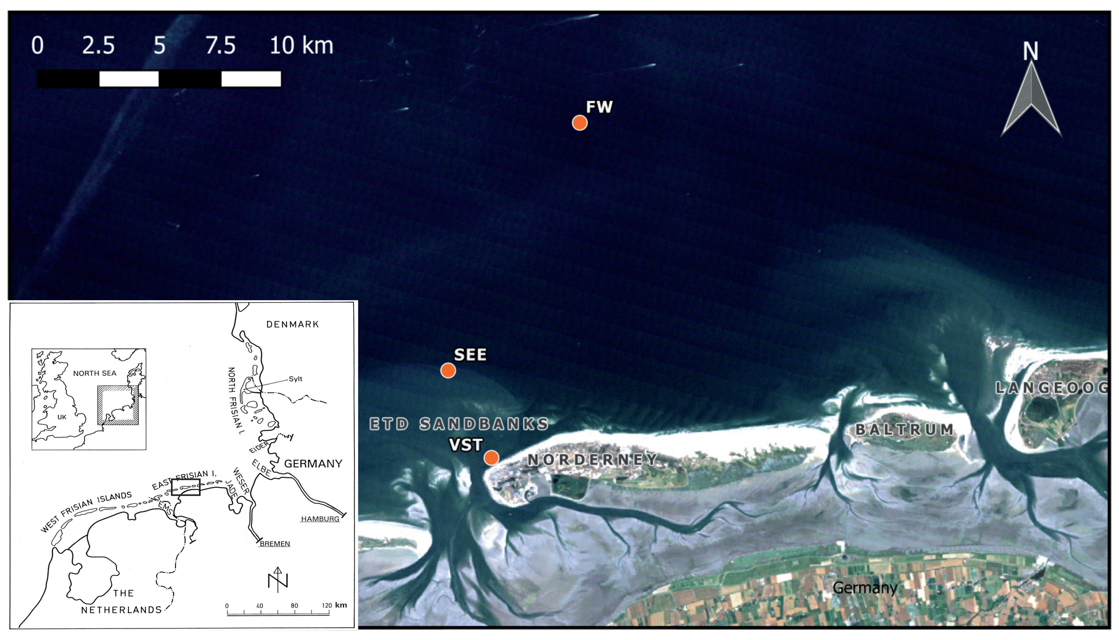

Figure 1: Nearshore area of Norderney with buoy positions.

Buoy Data

► Data available from 2004 to 2017 – with many missing values

Table 1: Significant wave height statistics.

Position	Min [m]	Max [m]	Mean [m]	Std. [m]	Number
FW	0.088	8.402	1.202	0.765	84 428
SEE	0.090	6.910	1.072	0.688	95 889
VST	0.038	3.009	0.626	0.393	88 805

Methodical Approach

► **Parameter Selection**
Input: significant wave height, wind speed/direction, water level
Output: significant wave height

► **Data Preprocessing**
Outlier Detection
Standardizing
One-Hot Encoding
PCA of the energy density spectrum

► **Model Development**
Choice of suitable statistical models
Choice of suitable model parameter (number of neurons, epochs, batchsize, etc.)

► **Model Training**
Calculation of optimal weights

► **Model Validation**
Analysis of RMSE and r of the respective model settings

► **Model Discussion**
Comparison of different models (LinReg, FFNN, LSTM)
Missing values
Parameter selection
Variability through parameters vs. variability through model selection

Model Setup

Table 2: Six different variations of forecast and reconstruction. A, B and C denote the significant wave height at position VST, SEE and FW, respectively. U denotes the wind speed, UDir the wind direction and W the water level. (1): Information of same location but different times. (2): Information of same time t but different locations.

Model (LinReg /FFNN /LSTM)		Input	Output	Comment
I	single reconstruction	*A(t) B(t)*	*C(t)*	(2)
II	single forecast	*C(t-k)…C(t-1) k=48h*	*C(t)…C(t+l), l=24h*	(1)
III	multiple timesteps reconstruction	*A(t-k)…A(t) B(t-k)…B(t) k=8h*	*C(t)*	(2)
IV	multiple timesteps climate data reconstruction	*A(t-k)…A(t) B(t-k)…B(t) U(t-k)…U(t) UDir(t-k)…UDir(t) W(t-k)…W(t), k=8h*	*C(t)*	(2)
V	complex forecast	*C(t-k)…C(t-1) U(t-k)…U(t+l) UDir(t-k)… UDir(t+l) W(t-k)…W(t+l) k=48h*	*C(t)…C(t+l), l=24h*	(1)
VI	climate data reconstruction	*U(t-k)…U(t) UDir(t-k)…UDir(t) W(t-k)…W(t), k=8h*	*C(t)*	-

(3) Results and Discussion

Forecast and Reconstruction of Significant Wave Height

► Handling of missing values has a considerable influence on model quality
► Best results obtained with LSTM in setup IV (RMSE = 0.145 [m], r = 0.983)

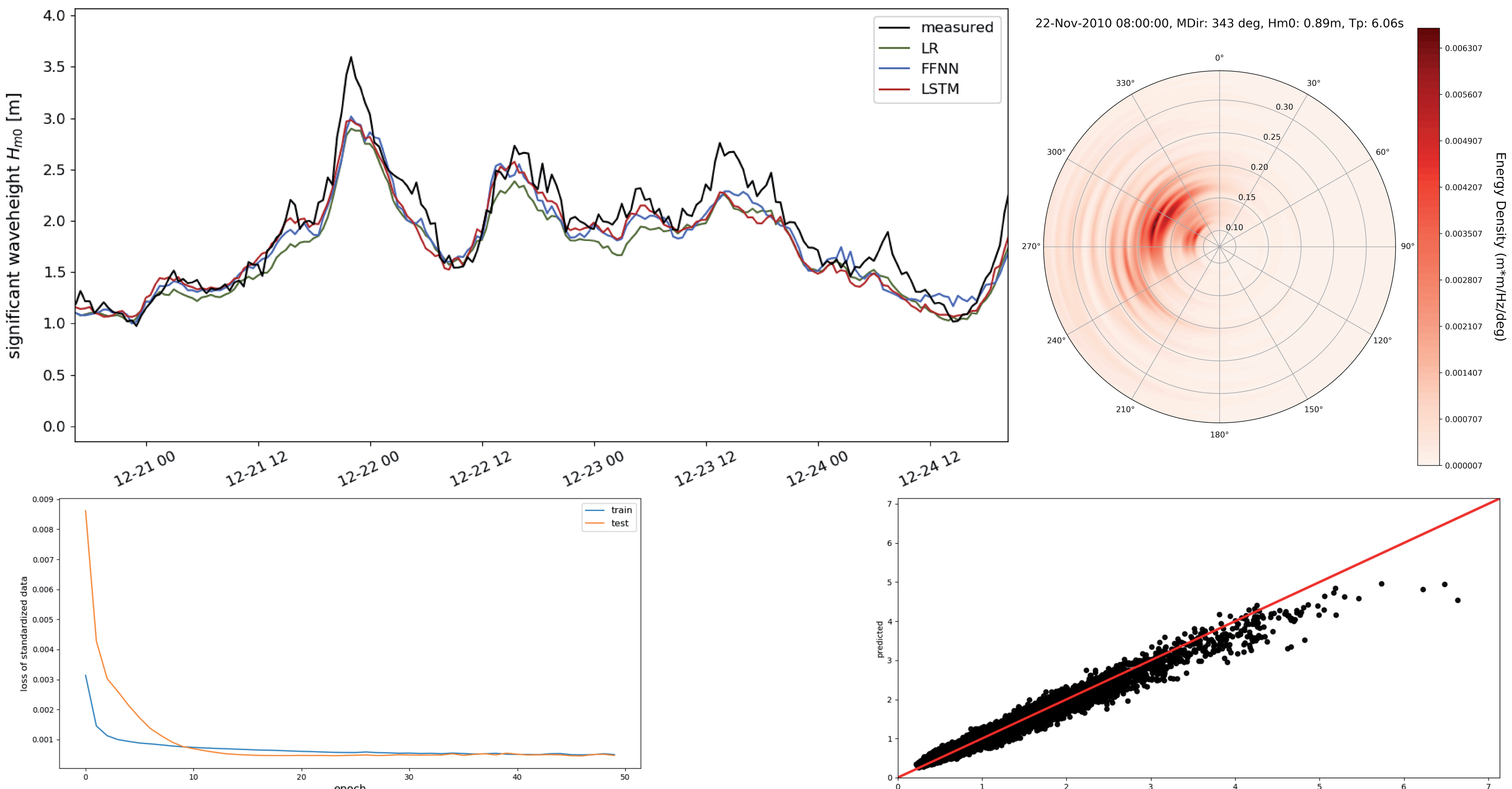

Figure 2: Top left: Reconstruction of significant wave height (model IV) for some test data. The LSTM, FFNN and LinReg are compared to the original measured data. Top right: Example of an energy density spectrum at position VST. Bottom left: Loss of the train and test data of model IV while the training process. Bottom right: Measured versus predicted significant wave height of LSTM in model IV.

Discussion of Different Models

► LSTM shows the best results for model setup I - IV compared to LinReg and FFNN
► LSTM shows bad results in setup V and satisfactory results in setup VI

Table 3: Comparison of all 18 model results of forecast and reconstruction of significant wave height.

experiment	method	available data	training data	test data	architecture ([neurons] batch size epochs)	test data RMSE	test data r
I	LinReg	37 554	27 010	10 544	-	0.202	0.965
	FFNN	37 554	27 010	10 544	10 64 50	0.194	0.967
	LSTM	37 554	27 010	10 544	5 64 50	**0.190**	**0.967**
II	LinReg	53 093	38 816	14 277	-	0.494	0.982
	FFNN	53 093	38 816	14 277	10 64 50	0.496	**0.986**
	LSTM	53 093	38 816	14 277	[49 49] 64 50	**0.493**	0.982
III	LinReg	32 729	22 358	10 371	-	0.183	0.969
	FFNN	32 729	22 358	10 371	10 64 50	0.177	0.970
	LSTM	32 729	22 358	10 371	[10 10] 64 50	**0.176**	**0.972**
IV	LinReg	32 722	22 358	10 364	-	0.151	0.980
	FFNN	32 722	22 358	10 364	10 64 50	0.146	0.980
	LSTM	32 722	22 358	10 364	[50 50] 32 50	**0.145**	**0.983**
V	LinReg	51 358	37 081	14 277	-	**0.281**	0.777
	FFNN	51 358	37 081	14 277	150 32 50	0.333	**0.784**
	LSTM	51 358	37 081	14 277	[25 15] 50 100	0.536	0.769
VI	LinReg	82 717	67 726	14 991	-	0.387	0.851
	FFNN	82 717	67 726	14 991	10 64 50	**0.362**	**0.873**
	LSTM	82 717	67 726	14 991	[10 10] 64 50	0.378	0.858

(4) Conclusions

General

- Complete exclusion of missing values led to the best results
- Model setting IV is proposed for the reconstruction of significant wave height time series
- Possibility to save costs by ending a buoy position and reconstructing the data from neighboring buoys
- Extreme events were consistently underestimated

LSTM specific

- First application of LSTM to forecast and reconstruct significant wave heights
- LSTM showed slightly better results than linear regression and FFNN, thus the LSTM is proposed
- Improving the way in which the energy density spectrum is pre-processed as input

References

Durán-Rosal, A.M., Hervás-Martínez, C., Tallón-Ballesteros, A.J., Martínez-Estudillo, A.C., Salcedo-Sanz, S., 2016. Massive missing data reconstruction in ocean buoys with evolutionary product unit neural networks. Ocean Engineering 117, 292–301. 10.1016/j.oceaneng.2016.03.053.

Kumar, N.K., Savitha, R., Al Mamun, A., 2017. Regional ocean wave height prediction using sequential learning neural networks. Ocean Engineering 129, 605–612. 10.1016/j.oceaneng.2016.10.033.

Qing, X., Niu, Y., 2018. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 148, 461–468. 10.1016/j.energy.2018.01.177.

Contact:
Christoph Jörges
joerges@uni-wuppertal.de

Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency - Coastal Research Station

NLWKN

BERGISCHE UNIVERSITÄT WUPPERTAL